# Spin : a 3-D Interface for Cooperative Work

Cédric Dumas, Samuel Degrande, Grégory Saugis, C. Chaillou, Patricia Plénacoste, Marie-Luce Viaud

# Spin : a 3-D Interface for Cooperative Work

Cédric Dumas (1,3)
dumas@lifl.fr

Samuel Degrande (1)
degrande@lifl.fr

Grégory Saugis (1)
saugis@lifl.fr

Christophe Chaillou (1)
chaillou@lifl.fr

Patricia Plénacoste (2)
plenacoste@univ-lille1.fr

Marie-Luce Viaud (3)
luce@ina.fr

(1) **LIFL** - Laboratoire d'Informatique Fondamentale de Lille
bâtiment M3 / Cité Scientifique / F-59 655 / Villeneuve d'Ascq cedex

(2) Laboratoire **Trigone** - CUEEP
bâtiment B6 / Cité scientifique / F- 59 655 / Villeneuve d'Ascq cedex

(3) **INA** - Institut National de l'Audiovisuel
4, avenue de l'Europe / F-94 366 / Bry sur Marne cedex

**Abstract**

In this paper, we present a three-dimensional user interface for synchronous co-operative work, Spin; which has been designed for multi-user synchronous real-time applications to be used in, for example, meetings and learning situations. Spin is based on a new metaphor of virtual workspace.

We have designed an interface, for an office environment, which recreates the three-dimensional elements needed during a meeting and increases the user's scope of interaction. In order to accomplish these objectives, animation and three-dimensional interaction in real time are used to enhance the feeling of collaboration within the three-dimensional workspace. Spin is designed to maintain a maximum amount of information visible. The workspace is created using artificial geometry —as opposed to true three-dimensional geometry— and spatial distortion, a technique which allows all documents and information to be displayed simultaneously while centering the user's focus of attention. Users interact with each other via their respective clones, which are three-dimensional representations displayed in each user's interface, and are animated with user action on shared documents. An appropriate object manipulation system (direct manipulation, 3D devices and specific interaction metaphors) is used to point out and manipulate 3D documents.

**Keywords:** Synchronous CSCW, CVE, avatar, clone, three-dimensional interface, 3D interaction.

## 1. Introduction

Technological progress has given us access to fields that previously only existed in our imaginations. Progress made in computers and in communication networks has benefited computer-supported cooperative work (CSCW), an area where many technical and human obstacles need to be overcome before to be considered as a valid tool. We need to bear in mind the difficulties inherent in cooperative work and in the user's ability to perceive a third dimension.

### 1.1 The shortcomings of two-dimensional interfaces

Current WIMP (Windows Icon Mouse Pointer) office interfaces have considerable ergonomic limitations [Schneiderman 1997].

*(a) Two-dimensional space does not display large amounts of data adequately:*

When it comes to displaying massive amounts of data, 2D displays have shortcomings such as window overlap and the need for iconic representation of information [Bly 1986]. Moreover, the simultaneous display of too many windows (the key symptom of Windowitis) can be stressful for users [Kahn 1995].

*(b) WIMP applications are indistinguishable from one another, leading to confusion:*

Window display systems, be they Xll or Windows, do not make the distinction between applications, consequently, information is displayed in identical windows regardless of the user's task.

*(c) 2D applications can't provide realistic representation:*

Until recently, network technology only allowed for asynchronous sessions (electronic mail for example); and because the hardware being used was not powerful enough, interfaces could only use two-dimensional representations of the workspace. Metaphors in this type of environment do not resemble the real space; consequently, it is difficult for the user to moving around within a simulated three-dimensional space.

*(d) 2D applications provide poor graphical user representations:*

As windows are indistinguishable and there is no graphical relation between windows, it is difficult to create a visual link between users or between a user and an object when the user's behavior is been displayed [Tanaka 1996].

*(e) 2D applications are not sufficiently immersive:*

Because 2D graphical interaction is not intuitive (proprioception is not exploited) users have difficulties getting and remaining involved in the task at hand.

## 1.2 Interfaces: New Scope

Spin is a new interface concept, based on real-time computer animation. Widespread use of 3D graphic cards for personal computers has made real-time animation possible on low-cost computers. The introduction of a new dimension (depth) changes the user's role within the interface, the use of animation is seamless and therefore lightens the user's cognitive load. With appropriate input devices, the user now has new ways of navigating in, interacting with and organizing his workspace.

Since 1995, IBM has been working on RealPlaces [Roberts 1998], a 3D interface project. It was developed to study the convergence between business applications and virtual reality. The user environment in RealPlaces is divided into two separate spaces:


- a "world view", a 3D model which stores and organizes documents through easy object interaction; and
- a "work plane", a 2D view of objects with detailed interaction, (what is used in most 2D interfaces).

RealPlaces allows for 3D organization of a large amount of objects. The user can navigate through them, and work on a document, which can be viewed and edited in a 2D application that is displayed in the foreground of the "world". It solves the problem of 2D documents in a 3D world, although there is still some overlapping of objects. RealPlaces does solve some of the problems common to 2D interfaces but it is not seamless. While it does introduces two different dimensions to show documents, the user still has difficulty establishing links between these two dimensions in cases where multi-user activity is being displayed..

In our interface, we try to correct the shortcomings of two-dimensional interfaces as IBM did in RealPlaces, and we go a step further, we put forward a solution for problems raised in multi-user cooperation. Spin integrates users into a virtual working place in a manner that imitates reality making cooperation through the use of 3D animation is possible.

Complex tasks and related data can be represented seamlessly allowing for a more immersive experience.

In this paper we discuss, in a first part, the various concepts inherent in simultaneous distant cooperative work (synchronous CSCW), representation and interaction within a three-dimensional interface. In a second part, we describe our own interface model and how the concepts behind it were developed. We conclude with a description of the various current and upcoming developments directly related to the prototype and to its assessment.

# 2. Concepts

When designing a three-dimensional interface, several fields need to be taken into consideration. We have already mentioned real-time computer animation and computer-supported cooperative work, which are the backbone of our project. There are also certain fields of the human sciences that have directly contributed to the development of Spin. Ergonomics [Bastien 1995], psychology [Carr 1995] and sociology [Nardi 1996] have broadened our knowledge of the way in which the user behaves within the interface, both as an individual and as a member of a group.

## 2.1 Synchronous Cooperative Work

The interface must support synchronous cooperative work. By this we mean that it must support applications where the users have to communicate in order to make decisions, exchange views or find solutions, as would be the case with teleconferencing or learning situations. The sense of *co-presence* is crucial, the user needs to have an immediate feeling that he is with other people; experiments such as *Hydra Units* [Buxton 1992] and *MAJIC* [Okada 1994] have allowed us to isolate some of the aspects which are essential to multi-media interactive meetings.

*Eye contact:*

A participant should be able to see that he is being looked at, and should be able to look at someone else.

*Gaze awareness:*

The user must be able to establish a participant's visual focus of attention.

*Facial expressions*:

These provide information concerning the participants' reactions, their acquiescence, their annoyance and so on.

*Gestures:*

Play an important role in pointing and in three-dimensional interfaces which use a determined set of gestures as commands, and are also used as a means of expressing emotion.

## 2.2 Group Activity

Speech is far from being the sole means of expression during verbal interaction [Cassel 1994]. Gestures (voluntary or involuntary) and facial expression contribute as much information as speech. Moreover, collaborative work entails the need to identify other people's points of view as well as their actions [Shu 1994] [Kuzuoka 1994]. This requires defining the metaphors which will enable users involved in collaborative work to understand what other users are doing and to interact with them. Researchers [Benford 1995] have defined various communication criteria for representing a user in a virtual environment. In DIVE (Distributed Interactive Virtual Environment), Benford and Fahlén lay down rules for each characteristic and apply them to their own system [Benford 1993].

They point out the advantages of using a clone (a realistic synthetic three-dimensional representation of a human) to represent the user. With a clone, eye contact (it is possible to guide the eye movements of a clone) as well as gestures and facial expressions can be controlled; this is more difficult to accomplish with video images.

In addition to having a clone, every user must have a telepointer, which is used to designate objects that can be seen on other users' displays.

### 2.3 Task-oriented Interaction

Users attending a meeting must be able to work on one or several shared documents. It is therefore preferable to place them in a central position in the user's field of vision, this increases her feeling of participation in a collaborative task. This concept, which consists of positioning the documents so as to focus user attention, was developed in the Xerox *Rooms* project [Henderson 1986]; the underlying principle is to prevent windows from overlapping or becoming too numerous. This is done by classifying them according to specific tasks and placing them in virtual offices so that a single window is displayed at any one (given) time.

The user needs to have an instance of the interface which is adapted to his role and the way he apprehends things. In a cooperative work context, the user is physically represented in the interface and has a position relative to the other members of the group.

### 2.4 The Conference Table Metaphor

Visually displaying the separation of tasks seems logical —an open and continuous space is not suitable. The concept of "room", in the visual and in the semantic sense, is frequently encountered in the literature. It is defined as a closed space that has been assigned a single task. A three-dimensional representation of this "room" is ideal because the user finds himself in a situation that he is familiar with, and the resulting interfaces are friendlier and more intuitive.

### 2.5 Perception and Support of Shared Awareness.

Some tasks entail focusing attention on a specific issue (when editing a text document) while others call for a more global view of the activity (during a discussion you need an overview of documents and actors). Over a given period, our attention shifts back and forth between these two types of activities [Harrison 1995]. CSCW requires each user to know what is being done, what is being changed, where and by whom. Consequently, the interface has to be able to support shared awareness. Ideally the user would be able to see everything going on in the room at all times (an everything visible situation). Nonetheless, there are limits to the amount of information that can be simultaneously displayed on a screen. Improvements can be made by drawing on and adopting certain aspects of human perception. Namely, a field of vision with a central zone where images are extremely clear, and a peripheral vision zone, where objects are not well defined, but where movement and other types of change can be perceived.

### 2.6 Interactive Computer Animation

Interactive computer animation allows for two things: first, the amount of information displayed can be increased, and secondly, only a small amount of this information can be made legible [Mackinlay 1991] [Robertson 1991]. The remainder of the information continues to be displayed but is less legible (the user only has a rough view of the contents). The use of specific 3D algorithms and interactive animation to display each object enables the user to visually analyze the data quickly and correctly.

The interface needs to be seamless. We want to avoid abstract breaks in the continuity of the scene, which would increase the user's cognitive load.

### 2.7 Navigation

We define navigation as changes in the user's point of view. With traditional virtual reality applications, navigation also includes movement in the 3D world. Interaction, on the other hand, refers to how the user acts in the scene: the user manipulates objects without changing his overall point of view of the scene.

Navigation and interaction are intrinsically linked; in order to interact with the interface the user has to be able to move within the interface. Unfortunately, the existence of a third dimension creates new problems with positioning and with user orientation; these need to be dealt with in order to avoid disorienting the user [Gomez 1994].

### 2.8 Manipulation

While navigation is restricted, the execution of an action is not. The direct manipulation of objects is more convenient and intuitive. The user can interact with the interface and manipulate objects directly. By enhancing the user's ability to move around within the interface we come closer to realistic interaction. We are working towards a model where representation, navigation and interaction are three-dimensional. The link between the virtual work environment and the real work environment are reinforced through the use of a 3D interface. In order to increase interaction, we want to create an environment where both hands can be used. Researchers [Kabbash 1994] have looked at the type of applications where bimanual interaction can be implemented without increasing the user's cognitive load. Kabbash explains that from his observations the use of two hands can be less productive than the use of one hand in cases where the application assigns an independent task to each hand. In certain cases, however, the use of two hands enables the user to adapt more quickly, to retrieve information faster, and to manipulate the interface with greater ease.

### 2.9 Deictic Gesturing

Increased scope for gestures also increases the incidence of problems related to hand position, namely, the perception of movement in real space and how it corresponds to movement in virtual space [Venolia 1993]. Interaction within the interface (mode) should correspond to the devices used to navigate (means). Research should look at the mode and the means of interaction concurrently.

With this in mind, we have decided to use 3D input devices [Fuchs 1995] [Dix 1998]. They can be put into three categories.

*Isometric Input Devices:*

Their resistance is infinite and they are stationary. They translate movement by measuring force and couple, so there is no direct correlation between what the hand does and what goes on in the interface. Another drawback is the lack of touch feedback (the user’s sense of proprioception is not exploited); this calls for extra adaptation time when performing complex tasks. Three D Trackballs are an example of this type of input device (figure 1a).

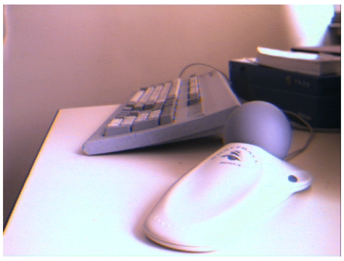


figure la: isometric device

figure 1b: isotonic device

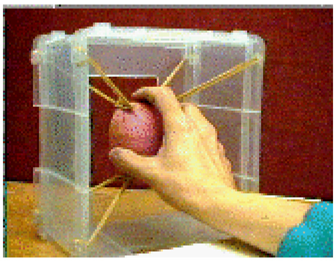


figure 1c: elastic device

*Isotonic Input Devices:*

Isotonic input devices move with the user and have no resistance. Their drawbacks are possible user fatigue following prolonged use. The data-glove or the acoustic devices are some examples of isotonic input devices (figure lb).

*Elastic Input Devices:*

Movement is translated by the amount of pressure applied**,** once released, they automatically return to their position in equilibrium. They are believed to correspond more to user proprioception and are thus easier to manipulate (figure 1c).

Opinions differ as to what type of input device obtains the best performance [Zhai 1994]. Isometric devices perform best for rate control (as is the case in robotics), whereas isotonic ones perform best for position control (in situations where there is a direct relation between hand and pointer movement).

# 3. Our Model

In this paper, we describe our interface model by expounding the aforementioned concepts, by defining spatial organization, and finally, by explaining how the user works and collaborates with others through the interface.

## 3.1 Spatial Organization

*The Workspace*

While certain aspects of our model are related to virtual reality, we have decided that since as our model is aimed at an office environment, the use of cumbersome helmets or gloves is not desirable. Our model's working environment is non-immersive. Frequently, immersive virtual reality environments lack precision and hinder perception: what human needs to perceive to believe in virtual worlds is out of reach of present simulation systems [Cadoz 1994]. We try to eliminate many of the gestures which are linked to natural constraints, (turning pages in a book, for example) and which are not necessary during a meeting. Our workspace has been designed to resolve navigation problems by reducing the number of superfluous gestures which slow down the user. In a real life situation, for example, people sitting around a table could not easily read the same document at the same time. To create a simple and convenient workspace, situations are analyzed and information which is not indispensable is discarded [Saugis 1997].

We often use interactive computer animation, but we do not abruptly suppress objects and create new icons; consequently, the user no longer has to strive to establish a mental link between two different representations of the same object. Because visual recognition decreases cognitive load, objects are seamlessly animated. We use animation to illustrate all changes in the working environment, i.e. the arrival of a new participant, the telepointer is always animated.

There are two basic objects in our workspace: the actors and the artefacts. The actors are representations of the remote users or of artificial assistants. The artefacts are the applications and the interaction tools.

*The Conference Table*

The metaphor used by the interface is the conference table. It corresponds to a single activity (our task-oriented interface solves the (b) shortcoming of 2D interface, see section 1.1). This activity is divided spatially and semantically into two parts. The first is a simulated panoramic view on which actors and shared applications are displayed. Secondly, within this view there is a workspace located near the center of the simulated panoramic screen, where the user can easily manipulate a specific document.


figure 2: objects placed around our virtual table

The actors and the shared applications (2D and 3D) are placed side by side around the table (figure 2), and in the interest of comfort, there is one document or actor per "wall". As many applications as desired may be placed in a semi-circle so that all of the applications remain visible. The user can adjust the screen so that the focus of her attention is in the center; this type of motion resembles head-turning. The workspace is seamless and intuitive, and simulates a real meeting where there are several people seated around a table. Participants joining the meeting and additional applications are on an equal footing with those already present. Our metaphor solves the (c) shortcoming of 2D interface (see section 1.1).

*Distortion*

If the number of objects around the table increases, they become too thin to be useful. To resolve this problem we have defined a focus-of-attention zone located in the center of the screen. Documents on either side of this zone are distorted (figure 3).

Distortion is symmetrical in relation to the coordinate frame x=0. Each object is uniformly scaled with the following formula:

$$x' = 1 - (1 - x)^{a} \quad , \; 0 \leq x \leq 1$$

where $\alpha$ is the deformation factor. When $\alpha=1$ the scene is not distorted. When $\alpha>l$, points are drawn closer to the edge; this results in centrally-positioned objects being stretched out, while those which are in the periphery are squeezed towards the edge. This distortion is similar to a fish-eye with only one dimension [Furnas 1986].

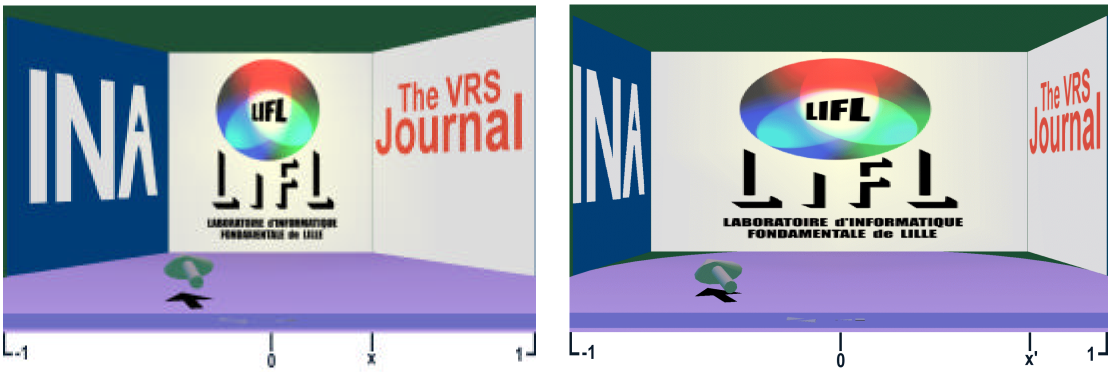



figure 3: Two examples of interface distortion

By placing the main document in the center of the screen and continuing to display all the other documents, our model simulates human field of vision (with a central zone and a peripheral zone). By reducing the space taken up by less important objects, an "everything perceivable" situation is obtained and, although objects on the periphery are neither legible nor clear, they are visible and all the information is available on the screen.

The number of actors and documents that it is possible to place around the table depends, for the most part, on screen resolution. Our project is designed for small meetings with four people for example (three clones) and a few documents (three for example). Under these conditions, if participants are using 17", 800 pixel screens all six objects are visible, and the system works.

*Everything Visible:*

With this type of distortion, the important applications remain entirely legible, while all others are still part of the environment. When the simulated panoramic screen is reoriented, what disappears on one side immediately reappears on the other. This allows the user to have all applications visible in the interface. In CSCW it is crucial that each and every actor and artefact taking part in a task are displayed on the screen (it solves the (a) shortcoming of 2D interface, see section 1.1).

*A Focus-of-Attention Area*

When the workspace is distorted in this fashion, the user intuitively places the application on which she is working in the center, in the focus of attention area. Clone head movements correspond to changes of the participants' focus of attention area. So, each participant sees the other participants' clones and is able to perceive their head movements. It gives users the impression of establishing eye contact and reinforces gaze awareness without the use of special devices. When a participant places a private document (one that is only visible on her own interface) in her focus in order to read it or modify it, her clone appears to be looking at the conference table (see clone on the figure 6).

In front of the simulated panoramic screen is the workspace where the user can place (and enlarge) the applications (2D or 3D) she is working on, she can edit or manipulate them. Navigation is therefore limited to rotating the screen and zooming in on the applications in the focus-of-attention zone.

## 3.2 Interaction

*The Pointer*

Our model uses bimanual interaction, so that the user has more interaction possibilities and because it simulates reality. The user has an input device which controls a pointer, the pointer's movements indicate changes in the user's hand position.

In a real life situation, users cannot go search for documents or tools without getting up from the table. With our room metaphor, the user does not have to navigate to select objects, he can manipulate them directly with the pointer, it can be moved throughout the entire meeting room. The scale difference between clone size (a fraction of the room volume) and user’s field of interaction has never been commented on by Spin users.

The isotonic input device used to control the pointer returns three dimensions of input data (x, y, and z). The interaction system is complemented by an isometric input device placed in the other hand, which is used for object manipulation and navigation, and for easy and immediate manipulation (rotation and zoom) of the simulated panoramic screen. The use of an isometric input device such as the 3D trackball in the non-dominant hand, should reduce hand movement and thus increase precision and decrease arm synchronization problems stemming from hand movement. With the addition of an isotonic input device (Polhemus™ trackers, for example) in the dominant hand, the interface is able to exploit the advantages of two sorts of devices (isotonic and isometric).

Translation of the dominant hand's movement in the "room" is immediately reflected in the interface by the pointer. Although the appropriate input devices are available to the user he may still lose his pointer when moving around in the interface. There are several ways of dealing with this problem. First, pointer orientation is used to indicate any change in direction and to enhance the impression of movement. Secondly, we use shading effects and the pointer's shadow is projected onto the floor. So as to maintain a constant size-intensity ratio, the shadow's intensity varies in relation to the cursor's distance from the floor. This helps the user to perceive meeting room depth accurately and to get his bearings quickly and easily [Kersten 1997]. Lastly, to make it easier to perceive the meeting room in 3D, we are able to place markers at regular intervals (a grid design, for example) on the interface conference table, this adds perspective. However, in one of our studies [Plenacoste 1998] -conducted to analyze the influence of different visual cues- we saw that artefacts other than shadows have a negligible influence on user depth perception in our interface. The study analyzed the effects of manipulating depth cues by measuring user accuracy in an aiming task in a three-dimensional computer environment. Subjects were asked to point at a specific target, a cube in a 3D room. For this task, we used two kinds of input devices: a 3D mouse and an isotonic device. We looked at several different possibilities: texture and no texture, pointer with shadow and without shadow, and target with shadow and without shadow. The results proved the superiority of using isotonic input devices and the usefulness of the shadow as a visual cue in a 3D computer environment. This selection system solves the (e) shortcoming of 2D interface (see section 1.1).

*3D Interaction*

The user should be able to interact in the interface by using the pointer. With the pointer, the user is able to select all of the graphical objects. Selecting an object must be simple. Our model uses visual cues to show that the object has been selected or that it can be selected: a graphical representation of a box appears progressively around the object. The closer the pointer is to the object, the more the surrounding box is visible and closer to the object. This progressive bounding box system greatly simplifies the manipulation of the pointer. Once an object has been selected, the commands which can be applied to it, appear around it as explicit symbols (*open*, *move*, etc.). Other options may appear depending on the application being used.

Once an object is selected, the user may want to manipulate it (figure 4). In order to maintain direct manipulation and to avoid widgets, we use an isometric device to rotate or move the object. This entails facilitating the work carried out on the interface by using the possibilities afforded by bimanual interaction whenever the application and the situation allow for it.

To access the objects’ functions, we currently use 3D circular menus with icons located around the object.

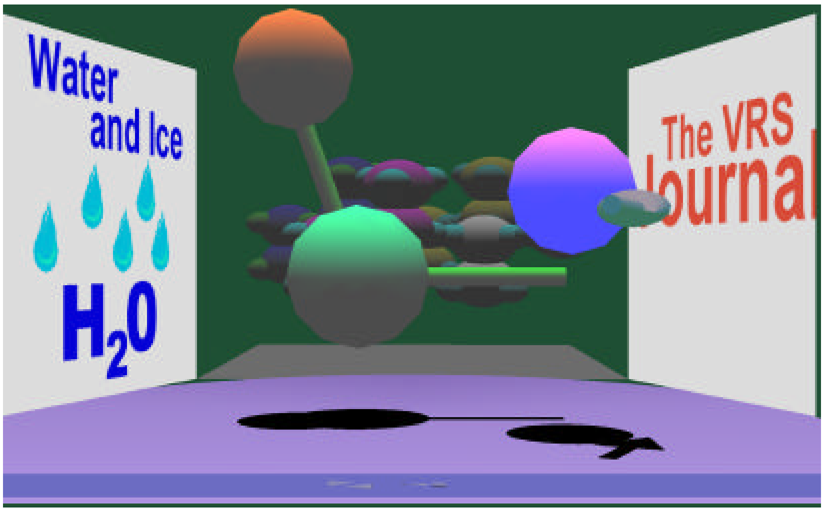



figure 4: selection/translation of an $H_2O$ molecule (only one user)

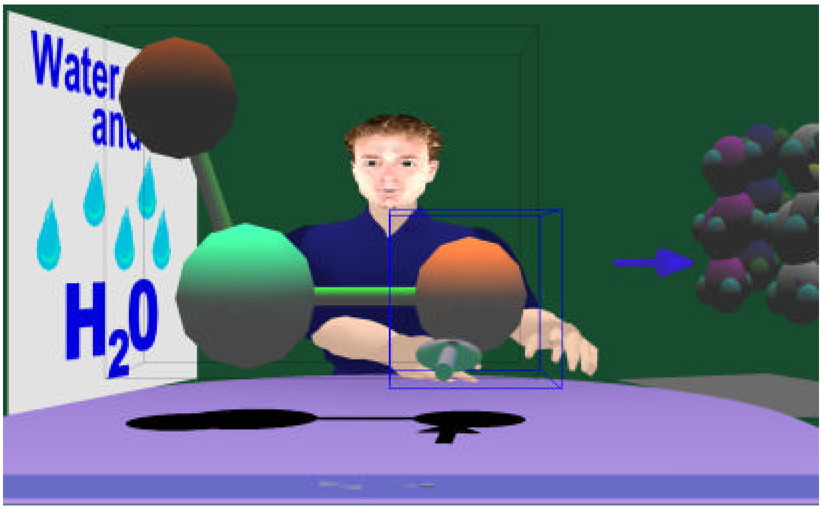



figure 5: clone of a distant user (two participants are connected)

*Actors and Artefacts*

There are many ways of representing the user; while video technology allows us to see participants, it requires high-performance networks. Moreover, the data are almost impossible to manipulate (e.g. it is impossible to illustrate where the participant is looking). Another problem is that camera angle has an impact on how users' gestures and attitudes are interpreted. This type of problem does not arise when clones are used as representations.

A clone is a synthesized three-dimensional representation of the user (figure 5). We use two photographs (front and side views) to obtain a three-dimensional model of the face as well as a complete texture of the head. These are then mapped onto a 3D model of the face. Our system has been developed by the researchers working on the Televirtuality project [Televirtuality]. This project has been elaborated to allow several physically distant persons to communicate in a virtual environment via their synthetic representatives The clone physically resembles the user and can, therefore, be used to identify him.

The clone's role is to show the remote participants' actions in an intuitive, coherent, and precise manner. We use clones because we can manipulate the clone’s entire body in order to illustrate participant behavior, head and eye movement can be controlled and can be used to indicate the focus-of-attention area, and the arms can be used for remote manipulation. It is relatively simple to place the avatar facing any given direction, or to place it in a resting position when animation data are not being transmitted through the network. Because the clone is a visual representation of the user and of his gestures, it can be used to transmit facial expressions, as well as head and arm movement [Viaud 1995]. A clone can even be used to elicit gestures during a conversation.

From a technical perspective, the use of clones implies that pictures are no longer transmitted in their "raw" video format; instead, only the data relevant to the clone's actions across time are transmitted. As there are fewer data (only a few bytes per analyzed image), narrow-bandwidth networks can be used for transmission. Clones, however, are not sufficient when it comes to identifying remote users’ finer actions. To deal with this, we have introduced a telepointer.

*The Telepointer*

A telepointer is a remote pointer which can be identified in a user's virtual workspace as being that of another user. Its position is determined by the user's pointer and its use is limited to shared applications. The telepointer's field of action is restricted, it is defined by shared application. The pointer's representation is directly related to that of its user (in the case of a clone it is an arm); its primary functions are designating objects and annotation; in these cases the arm can be used in nearby applications. When the telepointer is out of the shared object zone, it is seen on the other participants’ terminals in a stationary position next to its designated clone. 3D clones with their telepointers solve the (d) shortcoming of 2D interface (see section 1.1).

*Implementation*

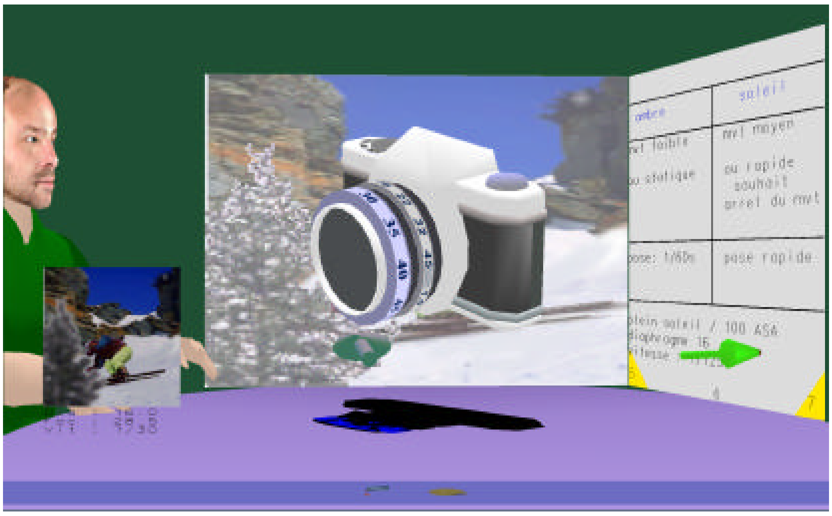
The pictures and diagrams which appear in this paper are from our interface prototype. This working model implements all of the concepts mentioned in this paper. The prototype was developed on a PC using Windows NT™. Our objective was to design our model using inexpensive equipment and for this reason, we intentionally avoided high-end equipment. As a graphical library, we used Open GL.

### 3.3 Example of a working situation in Spin

We have created a virtual camera simulator that will be used to teach participants how a real camera works. In order to asses the interface, three novice users will be asked to work together through Spin. They will share and manipulate a virtual camera and documents concerning the camera (figure 6). They will be able to take pictures with the simulator; all of the camera’s parameters will be taken into account: focus, shutter speed and aperture (figure 7). This situation will be compared to one where users will be asked to manipulate a real camera and communicate through a standard Internet communication tool (i.e. chat rooms and html pages). Results from these two experiments will be used by psychologists to analyze the specific aspects and possible advantages of synchronous co-operation and communication tools in Spin.


figure 6: two users, slides (the documents), and the camera on the table

figure 7: the camera with the real-time simulator (in front of the pointer) and the result



(the picture of the skier) on the wall in the center

## 4. Conclusion

We have presented Spin, a new 3D interface for synchronous CSCW, which is based on the concept that everything should remain visible at all times in a multi-user environment. The metaphor used in the interface is the conference table. The actors and the applications are positioned around the table without overlapping. The field of vision is divided into two sections, the central part of the screen is used as the focus of attention zone where everything is legible, while the areas on either side of the central zone are distorted. The interface uses a bimanual system, which allows users to easily select objects through a progressive bounding-box system. For 3D selection and manipulation tasks, Spin uses 2 three-dimensional input devices, isometric and isotonic, which are complementary.

The results from our first experiment proved that the current version of Spin is well suited for small meetings, where one or several documents are going to be explained and discussed in detail (i.e. presentations, research meetings, centred on a CAD object). With this in mind, we developed the camera simulator application to assess Spin. We want to be able to answer the following questions:

- Do users understand how 3D interaction works ? Do they use the different means of interaction available ?

• Do participants communicate in an efficient manner ? Are they able to learn how to use the camera ?

We observed that if standard lighting effects are used, the objects around the conference table lack volume (the whole scene has a short depth). In order to reinforce perception of third dimension, without modifying the geometry, visual effects such as textures, shadows and shading need to be used.

In the future, research needs to be oriented towards clone animation, and the amount of information clones can convey about participant activity. The aim being to increase user collaboration and strengthen the feeling of shared presence. New tools that enable participants to adopt an other participant's point of view or to work on an other participant's document, need to be introduced. Tools should allow for direct interaction with documents and users. We will continue to develop visual metaphors that will provide more information about shared documents, who is manipulating what, and who has the right to use which documents, etc.

In order to make Spin more flexible, it should integrate standards such as VRML 97, MPEG 4, and CORBA. And finally, Spin needs to be extended so that it can be used with bigger groups and more specifically in learning situation.

## Acknowledgements

We would like to thank Carmen Benito Garcia for her contribution to this paper.

The research reported here was supported by the National Center of Telecom Studies (CNET), the National Institute of Audiovisual (INA) and the regional council of Nord-Pas de Calais (France).

## References

Bastien, C. and Scapin, D.L. (1995). Evaluating a User Interface with Ergonomic Criteria. International Journal of Human-Computer Interaction 7: 105-121.

Benford, S. and Fahlén L. (1993). A Spatial Model of Interaction in Large Virtual Environments. ECSCW'93 Conference Proceedings.

Benford, S., Greenhalgh, C., Bowers, J., Snowdon, D. and Fahlén, L.E. (1995). User Embodiment in Collaborative Virtual Environments. CHI'95 Conference Proceedings: 242-248.

Bly, S. and Rosenberg, J. (1986). A Comparison of Tiled and Overlapping Windows. CHI'86 Conference Proceedings: 101-106.

Buxton, W.A.S. (1992). Telepresence: Integrating Shared Task and Person Spaces. Graphics Interface '92 Conference Proceedings: 123-129.

Cadoz, C. (1994). Le geste canal de communication homme/machine - la communication instrumentale. Technique et Science Informatiques 13(1): 31-61.

Carr, K. and England, R. (1995). Simulated and Virtual Realities: Elements of Perception. Taylor & Francis ltd., London.

Cassel, J., Pelachaud, C., Badler, N., Steedman, M., Achorn, B., Becket, T., Douville, B., Prevost, S. and Stone, M. (1994). Animated Conversation: Rule-based Generation of Facial Expression, Gesture and Spoken Intonation, for Multiple Conversational Agents. Siggraph'94 Conference Proceedings: 413-420.

Dix, J.A., Finlay, J.E., Abowd, G.D. and Beale, R. (1998). Human-Computer Interaction. second edition. Prentice Hall Europe.

Fuchs, P. (1995). Introduction aux Techniques de la Réalité Virtuelle. Ecole des Mines de Paris.

Furnas, G.W. (1986). Generalized fisheye views, Human Factors in Computing Systems. CHI’86 Conference Proceedings: 16-23.

Gomez, J.E., Venolia, D., van Dam, A., Fields, T. and Carey R. (1994). Why is 3d interaction so hard and what can we really do about it ?. Siggraph'94 Conference Proceedings: 492-493.

Harrison, B.L., Ishii, H., Vicente, K.J. and Buxton, W.A.S. (1995). Transparent Layered User Interfaces: An Evaluation of a Display Design to Enhance Focused and Divided Attention. CHI’95 Conference Proceedings: 317-324.

Henderson, D.A., Card, Jr. and Card, S.K. (1986). Rooms: the Use of Multiple Virtual Workspaces to Reduce Space Contention in a Window-Based Graphical User Interface. ACM Transactions on Graphics 5(3).

Kabbash, P., Buxton, W. and Sellen, A. (1994). Two-Handed Input in a Compound Task. CHI'94 Conference Proceedings.

Kahn, M.J. and Charnock, E. (1995). How to Prevent "Windowitis" in Your Graphical User Interface. ErgoCon'95 Conference Proceedings.

Kersten, D., Mamassian, P. and Knill, D.C. (1997). Moving cast shadows induce apparent motion in depth. Perception 26: 171-192.

Kuzuoka, H., Kosuge, T. and Tanaka, M. (1994). GestureCam, a Video Communication System for Sympathetic Remote Collaboration. Proceedings of ACM Conference on CSCW'94: 35-43.

Mackinlay, J.D., Robertson, G.G. and Card, S.K. (1991). The Perspective Wall: Detail and Context Smoothly Integrated. Proceedings of CHI'91: 173-180.

Nardi, B.A. (1996). Context and Consciousness. Activity Theory and Human-Computer Interaction. The MIT Press.

Okada, K., Maeda, F., Ichikawa, Y. and Matsushita, Y. (1994). Multiparty Videoconferencing at Virtual Social Distance: MAJIC design. Proceedings of ACM Conference on CSCW'94: 385-393.

Plénacoste, P., Demarey, C., Dumas, C. (1998). The role of static and dynamic shadows in a three-dimensional computer environment. Conference of AACE-Associatin for the Advancement of Computing in Education in coopération with WWW/internet bussinesses & Industry, WebNet 98, Orlando, USA.

Roberts, D. (1998). RealPlaces, 3D interface for Office Applications. IEE colloquium on The 3D Interface For The Information Worker 98(437).

Robertson, G.G., Card, S.K. and Mackinlay, J.D. (1991). Cone Trees: animated 3D visualisations of hierarchical information. CHI'91 Conference Proceedings.

Saugis, G., Dumas, C. and Chaillou, C. (1997). A new model of interface for synchronous CSCW. XIV Imeko World Congress, ISMCR 97 Topical Workshop on Virtual Reality and Advanced Man-Machine Interface, Tampere, Finland.

Schneiderman, B. (1997). Designing the User Interface: Strategies for Effective Human-Computer Interaction. Third edition. Addison-Wesley.

Shu, L. and Flowers, W. (1994). Teledesign: Groupware user experiments in three-dimensional computer-aided-design. Collaborative Computing 1(1): 1-14.

Tanaka, S., Okada, K., Kurihara, S. and Matsushita, Y. (1996). Desktop conferencing system using multiple still-pictures: Desktop-MAJIC. CSCW'96 Conference Proceedings.

Televirtuality. Televirtuality Project: http://www.ina.fr/Recherche/TV/TV.en.html

Venolia, D. (1993). Facile 3D Direct Manipulation. InterCHI'93 Conference Proceedings.

Viaud, M.L. and Saulnier, A. (1995). Real Time Analysis and Synthesis Chain. Proceedings of the 1995 International Workshop on Automatic Face and Gesture Recognition.

Zhai, S., Buxton, W. and Milgram, P. (1994). The “ Silk Cursor ”, Investigating Transparency for 3D Acquisition. CHI’94 Conference Proceedings.